*Chapter* 5

# Personalized Network Modeling in Psychopathology

**Abstract**

Two methodological trends have become prominent in the fields of psychopathology and psychiatry: (1) technological developments in collecting time-intensive, repeated, intra-individual measurements in order to capture symptom fluctuations and other time varying factors (e.g., emotions/affect) in daily life (i.e., time-series), and (2) an increasing number of statistical tools for for estimating assocations between these measurements (i.e., network structures) based on these time-series data. Combining these two trends allows for the estimation of intra-individual network structures. Using vector-autoregression (VAR), two networks can be obtained: a *temporal network*, in which one investigates if symptoms (or other relevant variables) predict one another over time, and a *contemporaneous network*, in which one investigates if symptoms predict one another in the same window of measurement. The network literature using these models has so far mostly focused on the temporal network. Here we argue that temporal relations between psychopathological variables might typically unfold within shorter time intervals (e.g., minutes) than the time intervals commonly and feasibly used in current time-series studies (e.g., hours). As a result, such temporal relations will be captured in the contemporaneous network, rather than in the temporal network. Both temporal and contemporaneous networks may highlight *potential* causal pathways—they are not definitive proof of causality but may lead to meaningful insights. As such, both types of networks function as *hypothesis generators*. We conclude the chapter with empirical examples of such analyses on symptom time-series data from clinical cases.

This chapter has been adapted from: Epskamp, S., van Borkulo, C.D., van der Veen, D.C., Servaas, M.N., Isvoranu, A.M., Riese, H., and Cramer, A.O.J. Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections.

## 5.1 Introduction

Recent years have witnessed an emergence of two distinct trends in the study of psychopathology. First, technological advances have permitted the gathering of intensive repeated measurements of patients and healthy controls with the Experience Sampling Method (ESM; Aan het Rot, Hogenelst, & Schoevers, 2012; Myin-Germeys et al., 2009; Wichers, Lothmann, Simons, Nicolson, & Peeters, 2012). With ESM, participants are measured repeatedly within short time intervals during daily life. For example, someone is queried five times a day during a period of two weeks on his or her level of insomnia, depressed mood, and fatigue since the previous measurement. We will term the time frame on which one reports the *window of measurement*. The resulting time-series data allow for the investigation of intra-individual processes (Hamaker, 2012). The second trend is the network perspective on psychopathology, in which mental disorders are interpreted as the consequence of a dynamical interplay between symptoms and other variables (Borsboom & Cramer, 2013; Cramer et al., 2010; Cramer & Borsboom, 2015). This literature uses network models in an attempt to understand and predict the dynamics of psychopathology. From this perspective, symptoms are not seen as passive indicators of a mental disorder but rather play an active role, making symptoms prime candidates for interventions (Borsboom, in press; Fried, Epskamp, et al., 2016).

Time-series data of a single individual offer a promising gateway into understanding the dynamical processes that may occur within that individual over time (e.g., Bringmann et al., 2013, 2015; Pe et al., 2015; Wigman et al., 2015). Such personalized network structures are typically estimated using a statistical technique called *vector-autoregression* (VAR; van der Krieke et al., 2015). Predominantly, VAR analyses have focused on the estimation of *temporal relationships* (relationships that occur between different windows of measurement). However, as will be outlined in Chapter 6, the residuals of the VAR model can be further used to estimate *contemporaneous relationships* (relationships that occur in the same window of measurement), which are not yet commonly used in the field. In this chapter, we argue that both network structures generate valuable hypothesis-generating information directly applicable to the study of psychopathology as well as to clinical practice. We focus the majority of the discussion on explaining contemporaneous partial correlation networks, as these are not yet often utilized in the literature of intra-individual analysis. We exemplify this by analyzing two ESM datasets obtained from patients.

## 5.2 Temporal and Contemporaneous Networks

In time-series data analysis with an average time-interval of a few hours, a typical default statistical assumption is violated: consecutive responses are not likely to be independent (e.g., someone who is tired between 9:00 and 11:00 is likely to still be tired between 11:00 and 13:00). The minimal method of coping with this violation of independence is the lag-1 VAR model (van der Krieke et al., 2015). In this model, a variable in a certain window of measurement is predicted by the same

variable in the previous window of measurement (autoregressive effects) and all other variables in the previous window of measurement (cross-lagged effects; Selig & Little, 2012)[1]. This model does not assume auto-correlations between larger differences in time (e.g., lag-2) are zero, but merely that such relationships can be fully explained by the lag-1 model. These autoregressive and cross-lagged effects can be estimated and visualized in a *network* (Bringmann et al., 2013). In this network, measured variables (such as symptoms) are represented by *nodes*. When one variable *predicts* another in the next window of measurement, we draw a *link* with an arrowhead pointing from one node to the other. We term this network the *temporal network*.

The predictive effects shown in the temporal network satisfy the assumption that in a causal relationship the cause must precede the effect. Therefore, these are often interpreted to be indicative of causal relationships. Only interpreting temporal coefficients, however, does not utilize VAR to its full potential. The residuals of the temporal VAR model are correlated; correlations in the *same* window of measurement remain that cannot be explained by the temporal effects. These correlations can be used to compute a network of *partial correlations* (Wild et al., 2010). In such a network, each variable is again represented by a node. Links (without arrowhead) between two nodes indicate the partial correlation obtained after controlling for both temporal effects and all other variables in the same window of measurement. We term this network the *contemporaneous network*[2].

Figure 5.1 shows an example of the two network structures obtained from a VAR analysis. These networks are estimated using ESM data of a clinical patient, and are further described and interpreted in Section 5.6. The temporal and contemporaneous networks were estimated at the same time, using the methodology outlined by Abegaz and Wit (2013). The temporal network (a) shows autoregressions (an arrow of a node pointing at itself) on three variables: 'tired', 'bodily discomfort' and 'concentration'. Thus, when this patient is tired she is likely still tired during the next window of measurement. There are cross-lagged relationships between several variables. For example, this patient being tired predicts her to ruminate more during the next window of measurement.. The contemporaneous network (b) shows, among other relationships, a relationship between 'relaxed' and 'nervous': when this patient was tired she was also more likely to relax poorer, *as reported during the same window of measurement*. This can be seen as the direct consequence of a plausible causal relationship: being nervous might lead you to feel less relaxed (or vice-versa). There is no reason why such a causal relationship should take a few hours to occur, which brings us to the main point of this chapter.

[1] VAR can be seen as an ordinary regression where the predictors are lagged variables.

[2] The contemporaneous network should not be confused with a network of lag-0 (partial) correlations. Such a network would (1) not take into account that responses are not independent, and (2) present a mixture of temporal and contemporaneous effects. Thus, we obtain the contemporaneous network from the residuals of the VAR model, since only then relationships *between* windows of measurement and relationships *within* windows of measurement are separated.

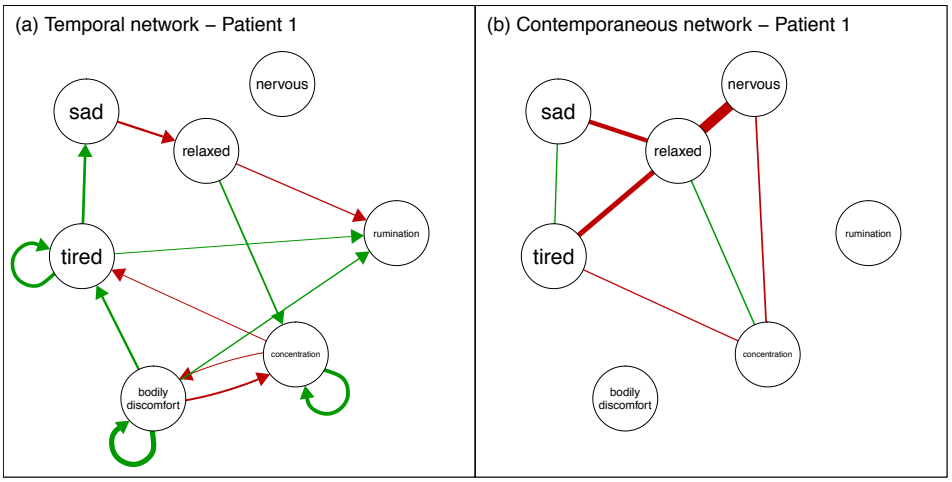


Figure 5.1: Two network structures that can be estimated with time-series data analysis, based on data of a clinical patient ($n = 52$) measured over a period of two weeks. The model was estimated using the *graphicalVAR* package for R. Circles (nodes) represent variables, such as symptoms, and connections (links, both undirected drawn as simple lines or directed drawn as an arrow) indicate predictive relationships. Green links indicate positive relationships, red links indicate negative relationships, and the width and saturation of a link indicates the strength of the relationship. The network on the left shows a *temporal network*, in which a link denotes that one variable predicts another variable in the next window of measurement. The network on the right shows a *contemporaneous network*, in which links indicate partial correlations between variables in the same window of measurement, after controlling for all other variables in the same window of measurement and all variables of the previous window of measurement.

## 5.3 Causation at the Contemporaneous Level

In a typical ESM study, the time between consecutive measurements is a few hours[3]. As such, the temporal network will only contain predictive effects of measured variables on other measured variables about a few hours later. However, it is likely that many causal relationships occur *much faster* than a timeframe of a few hours. Take for example a classical causal model:

$$\text{Turn on sprinklers} \rightarrow \text{Grass is wet.}$$

Turning on the sprinklers causes the grass to become wet. This causal effect occurs very fast: after turning on the sprinklers it takes perhaps a few seconds for the grass to become wet. If we take measures of sprinklers ("on" or "off") and the wetness of the grass every two hours, it would be rather improbable to capture the

[3]Notable exceptions are sampling designs in which individuals are asked to fill-out questionnaires once a day or week.

case in which the sprinklers were turned on *just* before the grass became wet. As such, the temporal network would *not* contain a connection between turning on the sprinklers and the grass being wet, and likely *would* only contain a temporal auto-regression of the grass being wet (because it takes time for grass to dry). However, after controlling for this auto-regression, we most likely would find a connection between these variables in the contemporaneous network: in windows of measurement where the sprinklers were on we would likely also find that the grass was wet.

We can think in a similar vein about psychopathological causal relationships. For example, a patient suffering from panic disorder might anticipate a panic attack by experiencing somatic arousal (e.g., sweating, increased heart rate):

$$\text{Somatic arousal} \rightarrow \text{anticipation of panic attack}$$

In this structure, an arrow indicates that whatever is on the left causes whatever is on the right. This patient anticipates a panic attack, because the patient is experiencing somatic arousal. This causal effect would likely occur within minutes, not hours. Someone who experiences somatic arousal between 13:00 and 15:00 might still experience somatic arousal between 15:00 and 17:00. Thus, we can expect to find auto-regressions. However, between somatic arousal and anticipation of panic attack we would likely only find a contemporaneous connection.

In sum, relations between symptoms and other variables can plausibly unfold faster than the time-window of measurement; such relationships will be captured in the contemporaneous network. Figure 5.1 showed, however, that the contemporaneous network has no direction (links have no arrow-heads). To understand *how* such undirected networks can still highlight potential causal pathways, we need to delve into the literature on estimation of networks in psychopathology.

## 5.4 Partial Correlation Networks

As outlined above, the contemporaneous relationships can be interpreted and drawn as a network of partial correlations. In this section, we describe how such partial correlation networks can be interpreted and how links in such a network can be seen as indicative of causal relationships. Partial correlation networks are part of a more general class of undirected (i.e., no arrows) networks (formally called Markov Random Fields; Lauritzen, 1996) that have been introduced to psychopathology in response to the call for conceptualizing psychopathological behavior (such as symptoms) as complex networks (Borsboom et al., 2011; Cramer et al., 2010). After the initial introduction of partial correlation networks to the psychopathological literature (Borsboom & Cramer, 2013), the use of undirected networks in psychopathology gained considerable traction following the introduction of easy-to-use estimation methods and publicly available software packages for both estimation and visualization (Epskamp et al., 2012; van Borkulo et al., 2014). Ever since, such network structures have been extensively applied to research in the fields of psychopathology and psychiatry, such as comorbidity (Boschloo et al., 2015), autism (Ruzzano, Borsboom, & Geurts, 2015), post-traumatic stress disorder (McNally et al., 2015), psychotic disorders (Isvoranu, van Borkulo, et al.,

2016; Isvoranu, Borsboom, et al., 2016), major depression (Fried et al., 2015; van Borkulo et al., 2015), and clinical care (Kroeze et al., 2016).

Partial correlation networks have become so prominent because they present a relatively easy method to estimate and visualize potential causal pathways, while taking into account that observational data (i.e., no experimental interventions) only contains limited information on such causal relationships. In observational data, causality is reflected only in the conditional independence structure (Pearl, 2000). Conditional independence means that two variables are no longer correlated at fixed levels of a third variable. A partial correlation network shows conditional independence, because when the partial correlation between two variables after conditioning on all others equals zero, then that means two variables are conditionally independent. Therefore, two nodes that are not directly connected via a link are conditionally independent.

Taking again the patient described above suffering from a panic disorder, suppose we expand the causal structure to include this patient's pathway related to avoiding feared situations:

Somatic arousal $\rightarrow$ anticipation of panic attack $\rightarrow$ avoidance of feared situations.

Anticipating a panic attack might cause this patient to avoid feared situations, such as malls or busy shopping streets[4]. The causal structure indicates that we would expect to be able to predict this patent avoiding feard situations given that he or she is experiencing somatic arousal. However, if we already know this person is anticipating a panic attack, we already predict that this person will avoid feared sitatations. Then, observing somatic arousal on top of anticipating the panic attack does *not* improve this prediction. Thus, we would expect non-zero partial correlations between somatic arousal and anticipation of panic attack, and between anticipation of panic attack and avoidance. We would furthermore expect a partial correlation of *zero* between somatic arousal and avoidance behavior; somatic arousal and avoidance behavior are conditionally independent given the anticipation of a panic attack. Consequently, we would expect the following partial correlation network:

Somatic arousal — anticipation of panic attack — avoidance behavior.

Finding such a partial correlation network often does not allow one to find the true direction of causation. This is due to two technical arguments: (1) equivalent models explain the same conditional independencies and (2) these models only work when we can assume the causal structure is acyclic (i.e., contains no feedback loops). Concerning the first argument, we can summarize the above causal structure as $A \rightarrow B \rightarrow C$, in which $A$ and $C$ are conditionally independent given $B$. This conditional independence, however, also holds for two other models: $A \leftarrow B \leftarrow C$ and $A \leftarrow B \rightarrow C$ (Pearl, 2000). In general, we cannot distinguish between these three models using only observational data. Adding more variables only increases this problem of potentially equivalent models, making it difficult to construct such a network only from observational data. Even

[4]These relationships should be taken as an example. The direction of such effects is still at topic of debate, and likely differs from patient to patient (Frijda, 1988).

when such a network can be constructed, we need to assume that the structure is not self-enforcing. That is, a variable cannot cause itself via some chain (e.g., $A \rightarrow B \rightarrow C \rightarrow A$). In psychopathology, however, this assumption likely does not hold (in our example above: anticipating a panic attack might cause more somatic arousal). As a result of these problems with directed structures when temporal or experimental information is lacking, undirected networks have been more successful in this emergent research field. In an undirected network, the observation that A and C are conditionally independent given B is represented by only one model: $A - B - C$ (Lauritzen, 1996).

To summarize, the contemporaneous network is an undirected network without arrowheads. This network shows a link when two variables are not conditionally independent given both all responses in the previous window of measurement and responses of all other variables in the current window of measurement. When two variables are conditionally independent, no link is drawn. If a causal relationship were present, we would expect such a link, and if a causal relationship were not present, we would not expect such a link. Therefore, the links in the contemporaneous network can be indicative of causal relationships. However, as finding the direction of such relationships is hard, we do not attempt to do so and keep the links in the contemporaneous network without direction.

## 5.5 Generating Causal Hypotheses

The connections in both the temporal and contemporaneous network cannot be interpreted as true causal relationships except under strong assumptions. The pathways shown can only be *indicative* of potential causal relationships. Such a pathway is a necessary condition for causality (we would expect such relationships when there is a true causal effect), but not sufficient (the relationship can also be spurious and due to, e.g., unobserved causes; Pearl, 2000). Therefore, these networks can be seen as *hypothesis generating*. To test for causality one needs to investigate what happens after experimentally changing one variable. If fatigue causes concentration problems, we would expect concentration levels to change after experimentally making someone fatigued. Experimentally changing concentration levels should, on the other hand, not influence fatigue. Such causal testing can only be done experimentally; it is hard to infer causality from observational data, no matter how often and intensive someone is measured and how intensive the sampling rate is.

In addition to generating hypotheses on causal links, both networks also generate hypotheses on which *nodes* are important. The importance of nodes in a network can be quantified with descriptive measures called *centrality measures* (Costantini, Epskamp, et al., 2015; Newman, 2010; Opsahl et al., 2010). A node with a high centrality is said to be 'central', indicating the node is well connected in the network. Such a central node may be a prime candidate for intervention, as targeting this node will influence the rest of the system. This is not only the case for central nodes in the temporal network, but also for central nodes in the contemporaneous network. Even when a node has no temporal connections, it can still carry a lot of information on subsequent measurements, purely by being

central in the contemporaneous network. For example, if $A$ predicts $B$ and $C$ in the same window of measurement (contemporaneous links), and $B$ and $C$ both predict themselves in the next window of measurement (autoregressions), then as a result $A$ is able to predict $B$ and $C$ in the next window of measurement, even though no cross-lagged relationships might be found in the temporal network.

While experimental intervention is needed to test causal hypotheses, such hypotheses on causal relationships and central nodes might be hard to verify in practice. For example, one cannot wait with forming treatment plans until after lengthy experimental designs have been tested on a clinical patient. In addition, in intensive treatments for example, multiple nodes are likely to be targeted simultaneously; the causal effect of one particular node is hard to test. Furhermore, it might not be known how certain symptoms can be treated at all (e.g., feelings of derealisation when the patient is suffering from a comorbid depresonalisation disorder, a disorder that is often concurrent with a panic disorder). Still, the obtained insights are useful: the personalized networks can be discussed with the patient and, when the patient recognizes the discovered relationships, help to generate hypotheses and choose interventions that target these nodes (Kroeze et al., 2016).

## 5.6 Clinical Example

To exemplify how the described symptom networks can be utilized in clinical practice, we analyzed ESM data obtained from two patients treated in a tertiary outpatient clinic in Groningen. Patient 1 was a female patient, aged 23, who received cognitive behavioral therapy (CBT) for a severe panic disorder and a depressive disorder secondary to the panic disorder. Her response rate was 74%. Patient 2 was a female patient, aged 53 suffering from major depressive disorder, in early partial remission after having received electroconvulsive therapy (ECT). Her response rate was 93%, and data collection started one day after her last ECT session.

### Methods

The patients received an extensive briefing plus written user instructions for the ESM measurements. Direct support was available 24/7. Patient data were gathered during normal daily life with an ESM tool developed for an ongoing epidemiological study. With our secured server system (RoQua; roqua.nl; Sytema & Van der Krieke, 2013), text messages with links to online questionnaires were sent to the patient's smartphone. All items could be answered on a 7-point Likert scale varying from '1=not at all' to '7=very much'. Measurement occasions were scheduled five times a day every three hours for two weeks (maximal number of possible measurement is 70), and took three to five minutes to complete. The timing of the measurements was adjusted to their individual daily rhythm with the last measurement timed 30 minutes before going to bed. Patients were instructed to fill-out the questionnaires as soon as possible after receiving the text message. The patient received a reminder after 30 minutes, and after 60 minutes the link

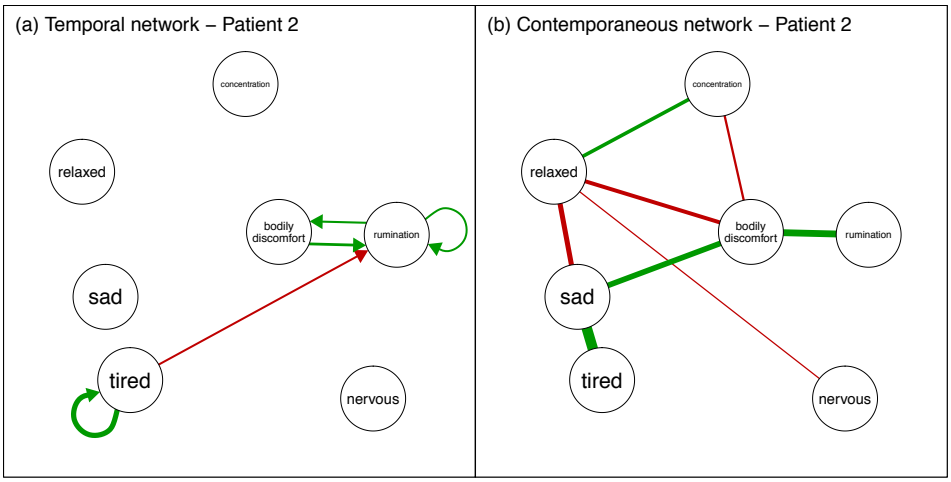


Figure 5.2: Temporal (panel a) and contemporaneous (panel b) network based on data of a clinical patient (n = 65) measured over a period of two weeks. The model was estimated using the graphicalVAR package for R.

was closed. The protocol used was submitted to the ethical review board of the University Medical Center Groningen, who confirmed that formal assessment was not required. Prior to participation, patients were fully informed about the study after which they gave written informed consent.

To estimate simple network structures with not many nodes, we selected seven of the administered variables that usually should interact with each other: feeling sad, being tired, ruminating, experiencing bodily discomfort, feeling nervous, feeling relaxed and being able to concentrate. Network structures were standardized as described by Wild et al. (2010) to avoid misleading parameter estimates in the network structure (Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016). The networks were estimated using the *graphicalVAR* package for R (Epskamp, 2015), which uses penalized maximum likelihood estimation to estimate model parameters (strength of connections) while simultaneously controlling for parsimony (which links are removed; Abegaz & Wit, 2013; Rothman et al., 2010). The graphicalVAR package estimates 2,500 different models, varying 50 levels of parsimony in the temporal network and 50 levels of parsimony in the contemporaneous network. Bayesian Information Criterion (BIC) model selection was used to select the best fitting model. A more detailed description of the estimation procedure is beyond the scope of this chapter. An introduction to model selection of regularized networks is provided in Chapter 2, and a methodological introduction to this modeling framework is provided in Chapter 6. We refer the reader to (Abegaz & Wit, 2013) for the estimation proceure used.

### Results

Figure 5.1 shows the two network structures of Patient 1. The temporal network in Panel (a) shows several connections involving bodily discomfort: whenever she experienced bodily discomfort, she ruminated more, felt more tired and was less able to concentrate three hours later. The contemporaneous network in Panel (b) shows that feeling relaxed plays a central role in the network. Whenever she was relaxed she experienced less sadness, tiredness, nervousness and was better able to concentrate (and vise versa, e.g., whenever she experienced less sadness she was more relaxed). In the case of Patient 1, therapy sessions revealed that intensively cleaning her house was her way of coping with stress. This lead to bodily discomfort and eventually rumination about her inability to do the things in the way she used to do things, resulting in a sad mood. Teaching her other ways to cope with stress broke this negative pattern.

Figure 5.2 shows the two network structures of Patient 2. The contemporaneous network in Panel (b) features more connections than the temporal network in Panel (a). In the contemporaneous network, the node bodily discomfort has a central role. Whenever Patient 2 experienced bodily discomfort (in her case, palpitations), she felt sadder, less relaxed, ruminated more and was less able to concentrate within the same window of measurement. This fits the pathology of a panic disorder where bodily sensations are interpreted catastrophically. The temporal network shows the effects over time and highlights a potential feedback loop, where bodily discomfort rumination (in her case, catastrophic interpretations of the bodily sensations) leads to more attention to bodily discomfort, causing more rumination. Feeling tired seems also to lead to more rumination in time.

## 5.7 Conclusion

In this chapter we argued that when analyzing intra-individual time-series data in clinical settings, researchers should focus on both temporal and contemporaneous relationships. While temporal networks are commonly estimated and interpreted in the network approach to psychopathology (e.g., Bringmann et al., 2013), contemporaneous networks, especially when drawn as a partial correlation network, are not commonly used. We have argued that both contemporaneous and temporal networks can highlight meaningful relationships, interpretable and useable by patients and clinicians in treatment, as well as present researchers with hypothesis generating exploratory results on potential causal relationships. Such personalized knowledge can be used for intervention selection (e.g., choosing which symptoms to treat), as well as generate testable hypotheses pertaining to the individual patient that can be used to perform experiments. In addition to temporal relationship, contemporaneous relationships are also important in discovering psychological dynamics, as such relationships can also occur at a much faster time scale than the typical lag interval used in ESM studies.

A main limitation of the VAR method is that, even when contemporaneous networks are estimated, the results depend on the lag-interval used. If the lag interval is too long, meaningful relationships might not be retrieved (e.g., some dynamics might occur between days or weeks rather than hours). Conversely, if

the relationship is too fast, and dissipates fast, it might also not be retrieved (e.g., if the effect of a relation dissipates after minutes, it might not be captured in a design that measures persons hourly or slower). The optimal lag-interval is often unknown, and can even differ between individuals and different variables. The lag-interval used is typically chosen in part due to practical reasons; it is not feasible for a patient to fill out a questionnaire often during a day (e.g., 20 times a day). The data gathering can also not take too long (e.g., more than two weeks), as the VAR model typically assumes people do not structurally change (Haslbeck & Waldorp, 2016a; Wichers et al., 2016). While effects that are slower than the lag-interval could be captured in a second temporal network (e.g., a network between days in addition to a network between measurements; de Haan-Rietdijk, Kuppens, & Hamaker, 2016), such methods require more observations.

The aim of this chapter is not to argue against interpreting temporal coefficients; both temporal and contemporaneous effects contain meaningful information on how the observed variables relate to one-another. Regardless, we strongly argue that the temporal and contemporaneous relationships should not be over-interpreted, as these merely highlight potential causal pathways. So what is the use then of contemporaneous and temporal networks if they do not allow for causal interpretation? We argue that, for an individual patient, it is hardly relevant if relationships in his/her data are causal or not. What matters is that both types of networks give the clinician as well as the patient a personalized, and visualized window into a patient's daily life. Moreover, this personalized window comes with a host of opportunities to arrive at tailor-made intervention strategies (e.g., treating central symptom of patient), and to monitor progress (e.g., will "deactivating" central symptom result in the deactivation of other symptoms?). Additionally, discussing the idiographic intricacies of networks with the patient offers ample opportunity for the patient to gain insight into his/her strengths and pitfalls and for reinforcing a sense of participation in one's own care. Personalized care is on everybody's agenda, and rightly so; given its benefits, so should network modeling of psychopathology data at the level of the individual be.